M/EEG-Based Bio-Markers to Predict the MCI and Alzheimer's Disease: A Review From the ML Perspective

Yang, S., Sanchez Bornot, J., Wong-Lin, K., & Prasad, G. (2019). M/EEG-Based Bio-Markers to Predict the MCI and Alzheimer's Disease: A Review From the ML Perspective. *IEEE Transactions on Biomedical Engineering*, *66*(10), 2924-2935. [8640065]. https://doi.org/10.1109/TBME.2019.2898871

Link to publication record in Ulster University Research Portal

Published in:
IEEE Transactions on Biomedical Engineering

Publication Status:
Published (in print/issue): 31/10/2019

DOI:
10.1109/TBME.2019.2898871

Document Version
Author Accepted version

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

Download date: 05/05/2023

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

M/EEG-based Bio-markers to predict the Mild Cognitive Impairment and Alzheimer's disease: A Review from the Machine Learning Perspective

Su Yang, Jose Miguel Sanchez Bornot, Kongfatt Wong-Lin, and Girijesh Prasad, Senior *Member, IEEE*

***Abstract*— This work reviews the state-of-the-art neuromarkers development for the prognosis of Alzheimer's diseases (AD) and mild cognitive impairment (MCI). The first part of this study is devoted to reviewing the recently emerged machine learning (ML) algorithms based on electroencephalography (EEG) and magnetoencephalography (MEG) modalities. In particular, the methods are categorized by different types of neuromarkers. The second part of the review is dedicated to a series of investigations that further highlight the differences between these two modalities. Firstly, several source reconstruction methods are reviewed and their source-level performances explored, followed by an objective comparison between EEG and MEG from multiple perspectives. Finally, a number of the most recent reports on classification of MCI/AD during resting state using EEG/MEG are documented to show the up-to-date performance for this well-recognized data collecting scenario. It is noticed that the MEG modality may be particularly effective in distinguishing between subjects with MCI and healthy controls, a high classification accuracy of more than 98% was reported recently; whereas the EEG seems to be performing well in classifying AD and healthy subjects, which also reached around 98% of the accuracy. A number of influential factors have also been raised and suggested for careful considerations while evaluating the ML-based diagnosis systems in the real-world scenarios.**

***Index Terms*— Alzheimer's disease; mild cognitive impairment; MEG; EEG; biomarkers; neuromarkers.**

I. Introduction

With the continuous development of health care provisions, the life expectancy of humans has witnessed an increasing trend in most of the world. Take UK for instance, during years 2014 to 2016, the life expectancy at birth was 79.2 years for males and 82.9 years for females, which underwent a slight increment (one month for male and two weeks for female) compared to the year range 2013 to 2015 [1]. With the facilitation of better health care over the years, additional increase of the life expectancy can be expected from the developing countries. However, the expanded life expectancy also leads to a higher chance of getting the ageing-related brain disorders that can have devastating effects on our daily life, some of which are even the direct cause of death. Undoubtedly, as one of the most sophisticated organs of the human beings, brain is the least well-explored part of the body: the reasons behind many brain malfunctions are still not quite clear, even after centuries of research [2]. Dementia, particularly the Alzheimer's disease (AD), is one of such brain malfunctions that can potentially lead to fatal consequences, if left without proper treatment in time. Over 9.9 million new cases of dementia are diagnosed each year worldwide, which implies one new case every 3.2 seconds. Of all the dementia cases, about 60%-70% are classified as AD [3]. Unfortunately, the cause of the AD is rather poorly understood, around 70% of the risks is believed to be genetic and many of which are related to genes directly [3, 4]. However, there are clear visual evidences observed in patients who have been diagnosed with AD: the size and shape of their brains tend to change drastically compared with the healthy brains [5]. It is therefore possible to detect and start the diagnosis process early, hence preventing or at least delaying the brain from evolving to the typical AD stage. As the prodromal/transitional process before AD, the confirmation of mild cognitive impairment (MCI) has become a critical factor to predict the occurrence of AD in the long run [6]. Although MCI is considered a pre-stage of AD and other dementias, the gradual cognitive decline may also be directly caused by other factors including depression, heart disease, diabetes, stroke, high blood pressure and cholesterol, or it may coexist with these comorbidities, as has been noted in [7]. For instance, it has been observed that the rate of conversion from MCI to AD can be 10% per year [8], which point to the importance of devising early neuro-rehabilitation and drug-therapy programs to treat early symptoms.

Moreover, with the age span increasing, there is an ever-increasing demand for professional doctors with necessary expertise for MCI/AD treatment. This makes it a necessity to deploy a fast and potentially more affordable program for the diagnosis of MCI and early AD. With the rapid development in the Machine Learning (ML) field, in particular the

This work was supported by the Northern Ireland Functional Brain Mapping Project Facility (1303/101154803), funded by Invest Northern Ireland and the University of Ulster.

Su Yang, Jose Miguel Sanchez Bornot, Kongfatt Wong-Lin, and Girijesh Prasad are with the Intelligent Systems Research Centre, School of Computing, Eng & Intel. Sys., Ulster University, Derry~Londonderry, Northern Ireland, UK. (correspondence e-mail: s.yang1@ulster.ac.uk).

Copyright (c) 2017 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

advancements in novel implementations for big data analysis [9, 10], a whole category of efficient and arguably more accurate ML approaches for the detection of early stage AD is now more viable than ever before.

In recent years, several reports have emerged in the community using algorithms based on ML to discover biomarkers that are related to brain anatomical or functional characteristics, drug-treatment efficacy, disease mechanisms or general malfunctions of the brain. In a typical ML approach, usually features are explored to discover most critical diverging patterns among MCI/AD and healthy control subjects: the decision-making process of such classification/recognition only takes up a few seconds, which has therefore become one of the most appealing advantages over any human doctor. Another advantage of the ML approach is the efficiency of an expert system: given enough data, the system can be trained to learn incredible amount of information in a rather short period of time (the system training typically may take minutes to a few days maximum); whereas to train a qualified neurologist/clinician often needs years of hard working and internships. The ML-supported program can potentially reduce the cost of diagnosis considerably.

A number of clinical modalities have been developed to study dynamical changes of brain activity and also have demonstrated critical application for detecting and monitoring changes in brain disorders. These techniques can be routinely integrated into the ML framework. Among the frequently used neuroimaging modalities, electroencephalography (EEG) and magnetoencephalography (MEG) have been used to evaluate brain functional and connectivity changes in AD/MCI detections. A number of documented reviews/surveys have addressed the diagnosis of Alzheimer's diseases and MCI using EEG/MEG from different angles. For example, the EEG and MEG based source connectivity analysis techniques focusing on AD were reviewed in [11]; using MEG as a putative biomarker for AD prediction/detection was reported in [12]; the MEG and its general analysis techniques are covered in [13]; a critical review specifically devoted to the EEG and ERP biomarkers of Alzheimer's disease was proposed in [14]; one recent survey paper introduced various preclinical biomarkers for identifying AD and MCI based on EEG/MEG in detail from the medical/clinical perspective [15].

Biomarkers can be broadly categorized into two major types: pathophysiological and topographical markers. The indicators for brain amyloidosis and tautopathy, such as amyloid tracer PET scans, T-tau and P-tau, are considered pathophysiological, whereas these changes are accompanied by neuronal and synaptic atrophy, leading to brain metabolic and connectivity cascades that can be quantified using topographical markers such as DTI, fMRI, FDG-PET and M/EEG [16]. It is worth mentioning that for topographical markers, the M/EEG-based neuromarkers are good indicators for neurophysiological evaluation of the disease status and its progression process, however, they may not be quite effective for diagnostic purposes. Unlike the deposition of Amyloid-beta (Aβ) 1-42 or phosphorylated Tau in the brain, M/EEG-based topographical neuromarkers do not directly reflect the pathophysiological characteristics of MCI and AD in the brain [16]. This work aims to address the topic from a machine learning point of view, and particularly, reviews the recent studies mainly on the topographical neuromarkers that focus on EEG/MEG recorded in a resting state (RS) scenario.

Several effective neuromarkers have been tested to predict the MCI and diagnose the AD [16, 17, 18, 19] . For example in the frequency domain, the power spectral density (PSD) and related methods have long been one of the most frequently used techniques to explore effective neuromarkers in EEG/MEG research [20, 21]. In the time domain, the analysis of functional connectivity has been especially useful in the study of unusual changes in region(s) of interest and its communication [23]. For the RS M/EEG scenario, PSD-related neuromarkers are useful to study cortical synchronization/desynchronization, whereas the functional cortical/deep source connectivity is often evaluated by markers such as coherence, lagged linear connectivity and other indexes [16].

Entropy is another popular statistical measure that have proven to be useful in neuromarkers research at the regional and functional connectivity level (e.g. Shannon spectral entropy [24], cross-approximate entropy [25] and dispersion entropy [26]). These methods are also often combined and implemented on top of other time-frequency analyses, such as the methods based on wavelet and empirical mode decomposition.

The proposed neuromarkers in the literature critically rely on the underlying used modality. For instance, FDG-PET's neuromarkers are associated with decrease/increase of metabolism in ROIs, possibly related to Alzheimer's progression, whereas MRI is commonly used to measure brain atrophy, or more specifically atrophy in critical ROIs such as entorhinal and hippocampal areas due to its role in memory processes [27]. Similarly, fMRI can be used to study decreased/increased brain functional activity that may be related to changes in metabolism or atrophy. Independently, Tau-PET and Aβ PET's neuromarkers can reveal those ROIs where tau and Aβ abnormal proteins accumulate with adverse effects such as synaptic and neuronal degeneration. Above modalities are excellent to identify affected ROIs due to its advantageous spatial properties, though they are much less relevant to study critical changes in brain oscillations.

Due to its excellent temporal resolution and intrinsic properties, EEG/MEG are much better positioned to study the brain dynamical changes, and many techniques can be directly borrowed from other areas to study the neuronal communication phenomena, such as Hidden Markov Model from signal processing [28] and deep learning from artificial intelligence [29]. However, MEG arguably possess certain advantages over other modalities: Compared to EEG, 1) the helmet of the MEG machine usually contains large number of sensors (a typical MEG machine could have more than 300 sensors), whereas EEG high density montages are limited to under 200 sensors with the burden of increased hours of preparation to reduce the electrodes impedance and guarantee the signal quality; 2) The magnetic field is mostly unaffected by the scalp which otherwise has negative filtering effects over the EEG signal; therefore source reconstruction analyses are

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

potentially more accurate for MEG data. As opposed to fMRI and PET, 3) MEG signals may offer real-time data analysis and monitoring of brain dynamical states, and thanks to the high density of the sensor montage, the reconstructed brain activity in the source space may achieve a spatial accuracy as good as that obtained with fMRI. In sum, MEG offers the highest temporal resolution, and therefore facilitates much richer statistical analysis that may exploit directly the critical information contained in neuronal oscillations.

In this paper, we review the state-of-the-art neuromarkers for detecting/classifying the AD and MCI. It is subsequently organized into three main sections. Section II is devoted to the review of numerous neuromarkers in diagnosing the AD and MCI, which is further divided into four subsections for specific topics on different types of features. In Section III, we focus on addressing the problem from a comparative point of view. The M/EEG reported findings at sensor- and source-level analysis are firstly contrasted, followed by a systematic analysis of EEG and MEG results on MCI/AD classification, with particular interest in discussion of the most relevant results on resting state M/EEG for MCI/AD classification. Conclusions and discussion are presented in Section IV.

II. Neuromarkers for MCI and AD Detection

The neuromarkers for MCI and AD detection/classification are reported in this section. Accurately diagnosing AD has been an active research topic for the past several decades; despite its vague cause, many effective neuromarkers were proposed in the literature to identify the presence of such brain mal-functioning. However, the development of the features within the machine learning scope is still underway, as the neuromarkers reported so far do not provide adequate specificity and sensitivity values for application in clinical practice. In addition, most of the reported neuromarkers are based on studies using a relatively small number of samples, and that critically affect the reproducibility of these studies. Even though in some reports, the cross-validation or other techniques only showed a small error, there is a risk of overfitting due to using the data obtained from small number of participants [28, 29]. The typical neuromarkers are discussed below under four broad categories.

2.1 Time-domain neuromarkers

One of the most intuitive ways for the neuromarker extraction is to find the informative features in the time domain directly. Mamani et al. [17] reported experiments using the grand average P450 event-related potential (ERP) as the feature for a three-class classification problem: 15 AD patients, 20 MCI subjects and 26 normal controls participated in the data collection experiments using a 32-channel EEG cap. Subjects were instructed to perform a series of visual-based working memory tasks. The independent component analysis was employed for automatic noise removal; the nonparametric Kruskal Wallis test is adopted to measure the cluster correction and 5% significance level were used to test the difference amongst the three classes; these classes were analysed using the conventional k-mean clustering algorithm. While performing the working memory tasks, it was found that the fronto-centro-parietal electrodes captured the most distinctive ERP signals for the three-class classification.

Yu et al. [20] proposed to use permutation dis-alignment index (PDI) to measure the coupling strength between EEG series in the time domain. Data from a 16-channel EEG system was collected to compute this feature; 14 right-handed patients with a diagnosis of AD and 14 age-matched healthy subjects participated in the experiment. Only the middle 10 minutes of recordings with high confidence were preserved for feature extraction; it was found that the value of PDI is inversely correlated with the strength of functional connectivity. Coupled with another neuromarker namely graph index complexity (GIC), the combined PDI-GIC neuromarker pair (weight ratio 2:1) achieved classification rate of 98.9%, which improved by 6.4% compared to using only the best single marker PDI. The recognition performance is quite encouraging; however, the number of subjects involved in the study is relatively low to generalize well across the AD population.

2.2 Frequency domain neuromarkers

As a complement to the time domain analysis, frequency analysis captures the features from the angle of maximising the energy/power information. With this approach, often the time information of the signal is entirely sacrificed. Therefore, often both the time and frequency domain features were extracted for better representation of the signal. Hornero et al. [32] presented a recognition system that could discriminate AD patients from healthy subjects by analysing the MEG background activities. A number of biomarkers were employed to test the effectiveness of classification, including the median frequency and spectral entropy in the frequency domain, approximate entropy and Lempel-Ziv complexity in the time domain. There were 41 elderly subjects involved in the experiment, 20 of which were AD patients and the rest were healthy controls. It is worth mentioning that the employed MEG system contained only 148 channels (comparable with some EEG systems); the signals were segmented into epochs of 10 seconds' length. The linear discriminant analysis (LDA) and a forward stepwise LDA with a leave-one-out cross-validation scheme were used for feature selection and classification. The best classification rate was achieved using the spectral median frequency with an accuracy of 75.6%. By further sequentially combining the second-best feature, i.e. approximate entropy, the recognition accuracy increased to 80.5% (80.0% sensitivity, 81.0% specificity). The frequency domain power spectral density as the biomarker played a critical role in their experiment, and it appears that by combining frequency domain and time domain features, the system performance could receive a boost.

Using EEG modality, power-based neuromarkers have also been used for MCI classification. Ye et al. [21] conducted a series of tasks to distinguish the healthy controls and MCI patients using a 64-channel EEG recording system. Their experiments included 22 participants, half of which were MCI subjects. The epoch length was 2.5 seconds during segmentation; the relative power ratio (the ratios are obtained by computing the individual band powers divided by the overall band power) was computed for each of these epochs. According

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

to their report, theta band is the main abnormal rhythm, whereas no significant differences were found between the relative alpha powers of two groups. It is also found that the MCI subjects in general have higher ratios and absolute power; in particular, the left temporal area was found as the most affected region in the brain for subjects with MCI.

Mazaheri et al. [33] investigated the topic of whether the subtle anomalies in EEG activity of MCI patients during a word comprehension task could provide the evidence of the conversion to AD. The research involved 25 amnestic MCI patients, a subset of whom developed AD within 3-years, and the data from 11 elderly controls were used for the comparative analysis. The EEG data from 19 to 32 channels was recorded at 250 Hz; time-frequency representations of power as the biomarkers were calculated for each trial (1 second prior to word onset, and 1.5 second after). It is well known that the sensor-level data analysis tends to be inaccurate due to the problems associated with volume conduction, i.e. the nearby electrodes pick up activities from the same sources hence make the received signals mixed. The method proposed in [33], instead of performing source reconstruction, attempted to circumvent this issue by focusing on trial-by-trial negative correlations [24, 33], given it is unlikely that a common source generates a simultaneous increase and decrease in power of different frequencies at distant electrode sites. They employed LDA and SVM to verify the hypothesis through classification performance. The best sensitivity and specificity of 80% and 95% respectively were achieved by SVM. In addition, a *d*-prime of 2.51 was reported to highlight the separation between the means of the signal and the noise distributions based on sensitivity index [36].

Poil et al. [37] proposed to combine multiple EEG-based neuromarkers into a diagnostic classification index; in this way the conversion from MCI to AD may be better predicted. In their work, the data from 86 patients were obtained; all of the subjects were initially diagnosed with MCI and during 2 years' period, 25 patients of which converted to AD. Together 35 different markers were extracted (including spatial, temporal and spectral features), six of which from both the temporal and spectral domains were selected for the prediction. The best single biomarker provided a sensitivity of 64% and specificity of 62%; by combining the six selected features, the performance was found increased to a sensitivity of 88% and specificity of 82%. Given the data was obtained from a 21-channel system (much less raw information were captured compared with caps with high sensor density), these results indicated a clear advantage of employing multi-biomarker analysis in the clinical field.

Recently, the cross-frequency coupling in cognition has become a new trend in neuromarker development. Dimitriadis et al. [38] investigated the phase coherence measure, in particular using the phase-to-amplitude estimator as a feature for the MCI vs. HC (healthy control) classification problem. Based on the EEG data captured from only a single sensor (Pz), a high recognition of 95% was achieved by using a poll of subjects containing 25 MCIs and 15 HCs. The data collection proceeded while subjects were performing the classic oddball tasks, with each trial of 1 second, in total 30 trials were obtained for each subject. Based on the leave-one-out cross-validation, their experimentations showed a high performance with 96% of sensitivity and 93% of the specificity.

2.3 Entropy and Complexity

Neuromarkers derived from the principle of entropy and complexity are found to be one of the most popular features for AD and MCI detection in the literature. Generally, researchers have been attempting to develop these kinds of neuromarkers broadly in two conventional ways: 1) entropy or entropy-related features computed using the time domain signals; 2) analysis of the complexity or entropy of the signals from the frequency domain perspective. The follow-up sub-sections are devoted to review a few of these researches from these two aspects.

2.3.1 Time-domain analysis

Gómez et al. [25] reported an accuracy of 70.83% (66.67% sensitivity, 75% specificity) using cross-approximate entropy to classify between AD patients and healthy controls. Cross-approximate entropy is a measure of asynchrony between two paired time-series [39]; five minutes of data were recorded using a 148-channel MEG system and the cross-approximate entropy of the signal between channels were computed. In total 24 subjects participated in the experiment, 12 of which were patients with AD and the rest were HCs. Compared with control subjects, a significantly higher synchrony was noticed between MEG signals from AD patients.

Azami et al. [26] did another investigation on the effectiveness of using different entropies as neuromarkers for AD analysis. A 148-channel whole-head magnetometer system was employed to collect the MEG data from a pool of 62 subjects, 36 of which were AD patients and the rest were the elderly controls. Four types of entropies, namely dispersion entropy (DisEn), fuzzy entropy, sample entropy (SampEn) and permutation entropy, were used to analyse the differences between the two classes. These features were directly computed in the time domain, using band-pass filtered signals (1.5-40 Hz) with segments of 10 seconds. The results indicated that the smallest p-value for AD patients vs. controls was obtained by using the DisEn; the computational efficiency of their newly proposed DisEn was also an appealing factor in the real-world scenario.

Another relatively straightforward method for AD detection (and arguably related to [25]) is to directly measure the complexity of the brain signals in the time domain. Gómez et al. [40] proposed to compute the Lempel-Ziv complexity of the MEG signals, which were recorded for 5 minutes in a relaxed state with a 148-channel whole-head magnetometer. Data from 10 patients with probable AD and 10 age-matched control subjects were used to test the effectiveness of the proposed biomarker, similar with the results reported in [39], it was found that the complexity level of the MEG signals from the AD class is significantly lower than the signals from the healthy controls.

The problem on the classification of AD and HC has also been addressed through the regularity and complexity measurements of the background activities in [41]. In their

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

study, five minutes of data was obtained using a 148-channel MEG system, 20 patients with AD and 21 healthy subjects contributed their data for the experiments. By employing the SampEn and multiscale entropy as neuromarkers, the MEG recordings from AD patients were found less complex and more regular than that from the HC subjects. The accuracies of 75.6% with SampEn, and 87.8% with multiscale entropy were reached.

2.3.2 Frequency-domain analysis

Using the same database as in [41] (20 AD and 21 HC), Poza et al. [42] computed the entropy-based markers in the frequency domain. Firstly, the relative spectral powers of the segmented MEG signals were computed, then a series of features including Shannon spectral entropy, Tsallis spectral entropy, generalized escort-Tsallis spectral entropy and Rényi spectral entropy were computed, the optimal markers were selected based on the results of Mann-Whitney *U* test. The classification was done using a binary logistic regression. The results suggested a significant decrease in irregularity of AD patients' MEG activity, which is in line with the observations that was reported in [41].

Entropy as a neuromarker can also be performed on top of other features for AD and MCI detections. For example, in [24], a series of entropies were computed to measure the classic frequency domain PSD. The objective was to classify between the MCI patients and the healthy controls using MEG data. Firstly, the PSDs from five typical brain regions were computed, these initial features were further used to compute Shannon spectral entropy, Tsallis spectral entropy and Rényi spectral entropy. Interestingly, in order to quantify the irregularity of MEGs, the Euclidean and the Wootters distances were employed as the disequilibrium measures. Using the data of a 10 seconds' recording with subjects in a relaxed state, awake and with eyes closed, a highest recognition rate of 64.3% were achieved using the Shannon spectral entropy as the neuromarker. A total of 18 patients with MCI and 24 healthy subjects participated in this experiment ([24]), which is comparable with the report from [32] in terms of subject numbers and employed modality (both were using MEG data for analysis). The experimental results (80.5% vs. 64.3%) seem to verify that the classification between MCI and HC subjects are more challenging than AD vs. HC classification.

Three-way classification of AD, MCI and HC subjects were also conducted by [43]. Data of 36 AD patients, 18 MCI subjects and 24 HC were recorded using a 148-channel whole-head magnetometer. PSD was computed as the initial feature, then the Jensen's divergence was adopted to measure the irregularity of the resulting PSDs. The results indicate significant changes of irregularity in the feature space for the data of AD patients, compared with MCI and heathy control subjects. The differences between MCI and HC are, however, less noticeable according to their report. These observations confirmed that the MCI is a transitional process towards AD stage from an ML point of view, and revealed that such classification is more challenging due to the prodromal nature of MCI.

In order to improve the recognition performance, the attempt of combing the time-frequency domain features and entropy-based neuromarkers were also made. Ruiz-Gómez et al. [44] tested a number of features, such as the relative power in the conventional frequency bands (i.e., delta, theta, alpha, beta, and gamma), median frequency, spectral entropy, sample entropy, and auto-mutual information. The experiments involved the classification for MCI, AD and HC. Each of the classes contained 37 subjects. Relevance and redundancy analyses were conducted to select the optimal set of features; LDA, quadratic discriminant analysis and multi-layer perceptron (MLP) were used to test and compare the classification performance. The results indicated MLP provided the highest performance for all the classification schemes: sensitivity of 82.35% and positive predictive value of 84.85% for HC vs. all classification task; specificity of 79.41% and negative predictive value of 84.38% for AD vs. all comparison.

Al-Nuaimi et al. [45] implemented a number of entropy and complexity based neuromarkers in the frequency domain. Data from 52 subjects were collected by a 19-channel EEG system, 20 of which were AD patients and the rest were healthy controls. Three types of neuromarkers namely Tsallis entropy (TsEn), Higuchi Fractal Dimension (HFD), and Lempel-Ziv complexity (LZC) were computed using data from the frequency bands. The results showed that AD patients have significantly lower TsEn, HFD, and LZC values for specific EEG frequency bands. In particular, the LZC provided the best overall classification performance with a sensitivity of 100%, specificity of 92.31% and the recognition accuracy of 95%. It is well-recognized that the classification between AD and HC is relatively easier than distinguishing HC and MCI; authors of the work in [45] also pointed out this challenge as their future work.

2.4 Other Neuromarkers

As the prodromal stage towards AD, MCI is found quite challenging to diagnose. In [46], a transform namely complete ensemble empirical mode decomposition (CEEMD) was employed to extract the neuromarkers from non-stationary MEG signals. A nonlinear dynamics measure based on permutation entropy was used as the neuromarker, which measures the characteristics of the resulting intrinsic mode functions after CEEMD. The analysis of variance (ANOVA) was used to select the computed entropy features, followed by an enhanced probabilistic neural network for the classification between normal and abnormal subjects. The data was collected using a 148 sensors MEG system; 18 MCI and 19 normal subjects participated in the experiment. A considerably high accuracy of 98.4% was achieved using the enhanced probabilistic neural network classifier. This seems to suggest that besides the conventional approaches, some less traditional algorithms, which are especially developed for the non-stationary signals can also be quite promising for AD/MCI detections.

Blind source separation methods is another viable approach that have been used to improve the diagnosis of AD patients. Data from 18 AD patients and 18 HC subjects were used for extracting a range of features [47]: mean frequency, spectral

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

entropy, approximate entropy, and Lempel-Ziv complexity. From the separated signals, the features with the most significant inter-class differences and least correlated were preserved based on the results from the Student's t-test. The preserved components were used to partially reconstructing the MEG channels, and the linear discriminant analysis was employed for classification. A significant boost in recognition rate from 72.2% to 80.6% was observed by performing the proposed blind source separation method.

A method based on the positive and unseen learning algorithm was implemented by Rasheed et al. for the identification of mild traumatic brain injury (mTBI), which was considered principally similar with diagnosing MCI using default mode network analysis in [48]. In detail, the classification was performed by devising default coherence limits between all pairs of MEG sensors for positive (control) group (7 subjects), and the assessment of severity (15 subjects) was carried out by using the positive and unseen learning method (single class model). The classification outcome of the proposed algorithm was compared with the original diagnosis: the average similarity between the ground truth and the algorithm performance was 79.52%, while the minimum similarity was 73% and the maximum similarity was 91%.

Another popular type of the biomarker is based on the cortical connectivity estimates. Gomez et al. [23] reported a diagnosis system to classify between MCI and HC subjects; data from 43 participants consisting of 18 MCI patients and 25 elderly controls were involved in the experiment. A 148-channel whole-head magnetometer (MAGNES 2500 WH, 4D Neuroimaging) was employed for the data collection, two connectivity measures, namely coherence and synchronization likelihood (SL) [49], were computed to measure the difference between the two classes. The results indicated that the coherence and SL mean values were lower in the MCI group than in control group at all frequency bands; it was also found the highest accuracy reached 69.8% in the beta band with both connectivity measures.

Researchers have been trying to combine the information of different frequency bands of the brain signals in order to improve the classification performance. Yu et al. [50] proposed an MEG-based system for the classification between AD and HC. In total 27 patients with AD and 26 HC subjects participated in the data collection, the MEG data were recorded using a 306-channel whole-head system. Several multiplex hub and heterogeneity metrics were computed to capture both overall importance of each brain area and heterogeneity of the connectivity patterns across frequency-specific layers. Their work indicated that the proposed multiplex brain networks analysis contains important information that cannot be revealed only by using frequency-specific brain networks. It was also found that MEG-based resting state multiplex networks in Alzheimer's disease were preferentially disrupted in hub regions, including regions in medial temporal lobe (left hippocampus), posterior default mode network and occipital regions.

The MAGIC-AD multicentre initiative was recently created to advance AD research, which has already produced MEG datasets of significance value including one dataset with 78 MCIs and 54 HCs in resting state, along with other two small datasets of 13 MCIs vs. 15 HCs and 11 MCIs vs. 13 HCs respectively, were employed [70]. In subsequent analysis by this group, the raw MEG signals were filtered into classic frequency bands (Theta (4-8Hz), Alpha (8-12 Hz), Beta (12-30 Hz), Gamma (30-45 Hz) and broadband (2-45 Hz)). The mutual information was used as the feature to estimate functional connectivity between all pairs of magnetometers. A series of experimental schemes with different database combinations were investigated to test the robustness of the proposed feature. The best classification performance was 83% of the accuracy, with 100% sensitivity and 69% specificity.

III. Comparative Review

The first part of this section is devoted to particularly reviewing the related source localization techniques on M/EEG; a number of factors on this technique that impact the classification are discussed. A comparative analysis of EEG and MEG in detecting AD and MCI is presented in the section 3.2, followed by a discussion on the pros and cons of these two modalities in the neurological field. Finally, a comparison of the recent works on MCI/AD/HC classifications in resting state is provided in section 3.3.

3.1 M/EEG-based MCI/AD detection at source-level

Despite numerous algorithms proposed for the source localization in literature, a great many challenges still remain: due to the ill-posed inverse problem, the selection of an appropriate inverse modelling algorithm becomes one of the most critical and debatable issues for the source-level analysis. Hincapié et al. [51] undertook a comparative analysis of some most popular source reconstruction methods, including the minimum norm estimate (MNE), linearly constrained minimum variance (LCMV) beamforming [52], and dynamic imaging of coherent sources (DICS) beamforming. A simulated sensor-level data through forward modelling based on a 275 channel CTF MEG system [53] configuration was used for testing the effectiveness of the algorithms. The MNE assumes a Gaussian distribution for the noise which is uncorrelated with the brain activity; despite the debatable observations, the main assumption of beamformer methods is that the oscillatory activities are uncorrelated between two different sources or dipoles [54]. These assumptions may be unrealistic and may have impact on the selection of these algorithms in the real-world scenarios. A few interesting observations were reported in [51]: though LCMV beamforming is a time-domain technique whereas the DICS beamforming works in the frequency domain, these two methods yielded similar performances in all the proposed test cases. Between the MNE and beamforming methods, it was found, for point-like sources (two coupled single-dipole sources), the spatial filters (LCMV and DICS) provided a better estimation of coherence, whereas MNE provided better coherence reconstructions when the simulated sources consisted of extended patches. These observations indicated the necessity of a scenario to imitate the real-life applications before deciding which source localization

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

method to use.

Another important factor for solving MEG/EEG inverse problems is the use of a template image or an individual MRI image for source reconstruction. A recent study conducted by Douw et al. [55] investigated the topic using data from 17 healthy participants who provided their MEG and MRI scans during a resting state recording with eyes closed. Relative power from six typical frequency bands for each region of interest after averaging were used as the neuromarkers. Functional connectivity (phase lag index) between each pair of regions was also calculated. It was found that there was no (systematic) bias or inconsistency between the results for the template and native MRI implementations, which is an important result in the field.

The combination of MEG and other modalities for feature extraction demonstrated good performance. Nakamura et al. [56] investigated the prodromal stages of AD based on MEG modality. A linearly constrained minimum variance beamforming technique was employed to perform the source reconstruction. MEG data obtained from 28 individuals with mild cognitive impairment and 38 cognitively normal individuals were used for feature extraction. The preliminary feature extracted was the regional spectral patterns, and then the integrating information from Pittsburgh compound B-PET, fluorodeoxyglucose-PET, structural MRI, and cognitive tests were analysed. The results demonstrated that regional spectral patterns of resting state activity could be separated into several types of MEG signatures [56], which may be used as useful biomarkers for the pre-dementia stages of Alzheimer's disease.

Usage of the source localization algorithm on EEG for the diagnosis of early AD was studied by Aghajani et al. [57]. Data obtained from 17 HC subjects and 17 subjects with AD were used for the data analysis. The source localization method was the standardized low-resolution brain electromagnetic tomography (sLORETA), the relative logarithmic power spectral density values from four conventional EEG bands (alpha, beta, delta, and theta) were extracted from 12 selected cortical regions. The results showed that the right temporal region reflected a significant difference between the two groups in all frequency bands; in the left brain hemisphere the theta band power increased whereas the alpha band decreased for AD patients. The classification performances using a support vector machine between AD and HC groups were accuracy of 84.4%, sensitivity 75.0%, and specificity of 93.7%.

Dimitriadis et al. [58] recently proposed a source-level analysis based on the reconstructed MEG signals. It investigated the performance of different analytic strategies of single-layer and multi-layer representations of functional brain networks. Three connectivity estimators namely phase locking value (PLV), the imaginary part of PLV (iPLV) and the correlation of the envelope (CorrEnv) were computed and used as the neuromarkers. Four minutes of resting state activity were obtained using a 306-channel Elekta Vectorview system from 24 MCI patients and 30 healthy controls. The source reconstruction was performed using an LCMV beamformer. Particularly, following the Yu et al.'s research [50], Dimitriadis et. al. [58] also proposed studying the intra and cross-frequency coupling or functional connectivity estimators as novel neuromarkers. They reported a highest classification accuracy of 98% using the CorrEnv feature. Despite the remarkable performance, the authors also pointed out the importance of testing the performance in a second blind cohort; due to the non-stationary characteristics of the time-series, using cross-validation within one database often tend to yield biased performance.

A recent investigation showed that the brain networks tend to facilitate information propagation across different frequencies, which was demonstrated by the analysis of multi-participation coefficients (MPC) [59]. In this work, the PSDs were used as features for the classification problem of AD vs. HC; each class contained the data obtained from 25 subjects. The data was collected using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX MEG system). To solve the inverse problem, a weighted Minimum Norm Estimate algorithm with overlapping spheres was employed. It was found that the regional connectivity in AD subjects abnormally distributed across frequency bands as compared to controls, causing significant decreases of MPC, which was similar with the trend found in the entropy-based analysis results. The best classification accuracy of 78.39% was reported for the proposed detection system.

Medvedeva and Yahno [60] reported an investigation based on using the EEG signals for the analysis of AD and MCI. Data from 131 AD, 45 MCI and 45 HC subjects were collected using a digital 19-channel scalp EEG device based on the International 10-20 system [61]. For each subject, 40 seconds of artefact-free EEG recording was kept and segmented into a window size of 2 seconds. Coherence measurements were computed and used as the biomarkers, combined with the eLORETA for the source-level analysis. Statistically significant differences between AD and MCI patients for theta band coherences were found. In addition, MCI subjects showed reduced coherence compared with healthy controls in certain regions.

Babiloni et al. [62] investigated the individual alpha frequency peak (IAF) and transition frequency (TF) after conventional FFT-based power spectral analysis, as features for the classification between MCIs and HCs. A 19-sensor EEG cap was used to capture the raw signals; eight bands (delta, theta, alpha 1, alpha 2, alpha 3, beta 1, beta 2, and gamma), and five ROIs (frontal, central, parietal, occipital, and temporal) were taken into the consideration during the data analysis. A source reconstruction method namely eLORETA was used to estimate the functional lagged linear connectivity solutions. It was reported that the best recognition performance for MCI vs. HC were sensitivity of 73% and specificity of 64%, led an accuracy of 68.5% and an area under ROC curve of 0.71.

An interesting investigation was conducted by Pineda-Pardo et al. [63] on the classification of the HC, single-domain MCI and multiple-domain MCI subjects. According to the report, Single-domain MCI (sdMCI) showed isolated memory impairment, whereas multiple-domain MCI (mdMCI) showed memory deficit accompanied by various degrees of impairment

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

Table 1. Some reports on MCI/AD detection using source reconstruction techniques

Modality	Neuromarkers	Source Reconstruction	Subjects	Classification Accuracy
MEG	Regional spectral patterns, integrating information from Pittsburgh compound B-PET, fluorodeoxyglucose-PET, structural MRI, and cognitive tests [56]	Minimum variance beamforming	28 MCI, 38 HC	Within HC to predict amyloid-β positivity: 76.3% Within MCI to predict amyloid-β positivity: 78.6%
EEG	Relative logarithmic power spectral density values [57]	sLORETA	17 AD, 17 HC	84.4%
MEG	Phase locking value (PLV), the imaginary part of PLV and the correlation of the envelope [58]	Linearly Constrained Minimum Variance beamformer	24 MCI, 30 HC	98% for the correlation of the envelope and 94% for the imaginary part of PLV
MEG	PSD [59]	Weighted Minimum Norm Estimate	25 AD, 25 HC	78.39%
EEG	coherence measurements [60]	eLORETA	131 AD, 45 MCI, 45 HC	N/A
EEG	Individual alpha frequency peak, Transition frequency [62]	eLORETA	75 MCI, 75 HC	68.5%
MEG	time domain connectivity matrix [63]	Minimum variance beamformer	22 sd MCIs, 30 md MCIs, 29	HC vs. sdMCI 86.27% HC vs. mdMCI 81.36% sdMCI vs mdMCI 84.62%

in cognitive domains such as executive functions, visual- spatial skills, and/or language. A 306-channel MEG system was employed for data collection, a database containing 29 HCs, 22 sdMCIs and 30 mdMCIs in resting state were used for this classification problem; an atlas-based anatomical parcellation of 66 regions was obtained for each subject. A minimum variance beamformer algorithm was adopted for the source reconstruction. The reconstructed time-domain connectivity matrix was directly used as features for classification. A number of classifiers, namely k-NN, LDA, SVM with different kernels were selected for distinguishing the three classes. The best accuracies were as follows: HC vs sdMCI 86.27%, HC vs mdMCI 81.36% and sdMCI vs mdMCI 84.62%.

Some recent works on the classification between MCI/AD and HC in the source-level are illustrated in Table 1. It appears: 1) different type of source reconstruction methods do not yield significant difference in performance if the database size is comparable; 2) while the number of subjects increases, the recognition performance tends to degrade drastically (98% compare to less than 70%). Here the modality information is purposely ignored, as it is hard to make a fair judgement given that very limited results are reported to date. In the next section, a further comparison is made to address the effectiveness of EEG and MEG on MCI/AD detection.

3.2 EEG vs. MEG: A comparative analysis

EEG as an electrophysiological brain monitoring approach has almost 100 years of usage history: the first human EEG recording was obtained by Hans Berger in 1924 [64]; since then EEG has been implemented in a series of different scenarios [64，65，66]. As one of the non-intrusive modalities, EEG not only benefits from its relatively easy deployment, but also it is considerably cheaper than MEG in equipment purchasing and maintenance. It is therefore popular among researchers for quantitative analysis of brain activations.

On the other hand, MEG received increasing attentions in recent years despite the need of high financial investment to prepare the related equipment for data collection and analysis. One of the major advantages of the MEG over EEG is that the magnetic fields are less distorted than the electric fields by the skull and scalp, which results in a better spatiotemporal resolution compared with EEG. High spatial resolution used to be an appealing advantage of image-based modality such as fMRI, but nowadays we can also reach a comparable quality of spatial resolution using MEG. Not least important, the preparation time for participants in the MEG system is much less than in EEG system (a few minutes vs. roughly one hour), which have a critical effect on the disposition of the participants and hence the data quality. A brief comparative analysis between EEG and MEG based on the state-of-the-art literature in relation to the AD and MCI detection is outlined below.

Table 2 and Table 3 illustrate some of the most recent research outcomes, separately for studies using EEG or MEG modality. For the sake of clarity and straightforward comparison, only the classification accuracy is reported here. It is noticed that classification accuracies are similar between analyses using these techniques, though slightly better results seem to be obtained using MEG modality particularly for HC vs. MCI scenario. We attribute the latter to the high density of MEG sensor arrays in contrast to EEG, and the simpler relation between neuronal activities generated inside the brain and externally collected signals. Otherwise, we observed that the classification between AD and HC has been found to provide better performance than the MCI vs. HC for both M/EEG, but this observation may be difficult to corroborate as there are only few studies simultaneously analysing all these 3 categories. As MCI is considered a prodromal stage of AD, it would be

Table 2. Comparison between two classification scenarios using EEG

EEG-based Report Accuracies	MCI vs. HC	AD vs. HC
Poil et al., 2013 [37]	85%	
Aghajani et al., 2013 [57]		84.4%
Al-Nuaimi et al., 2016 [45]		96%
Mazaheri et al., 2017 [33]	87.5%	
Ruiz-Gómez et al., 2018 [44]		82%
Yu et al., 2018 [20]		98.9%

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

Table 3. Comparison between two classification scenarios using MEG

MEG-based Report Accuracies	MCI vs. HC	AD vs. HC
Escudero et al., 2007 [47]		80.6%
Gómez et al., 2007 [41]		87.8%
Poza et al., 2007 [42]		85.4%
Hornero et al., 2008 [32]		80.5%
Bruna et al., 2010 [24]	64.3%	
Gómez et al., 2012 [25]		70.83%
Sanchez et al., 2016 [46]	98.4%	
Guillon et al., 2017 [59]		78.39%
Rasheed et al., 2017 [48]	79.52%	
Hernandez et al., 2018 [23]	69.8%	
Nakamura et al., 2018 [56]	78.6%	
Dimitriadis et al., 2018 [58]	98%	

interesting to study: 1) which neuromarkers appear consistently in both AD vs. HC and MCI vs. HC analyses, as well as which ones are less stable. 2) regarding the individual neuromarker trends, whether these tend to show ceiling effects, further deteriorate or show less impact in AD as compared to MCI, when considering HC as a baseline state. The results in Table 3 seem to suggest that the MEG have a great potential in detecting early signs of AD; a few rather high accuracies were reported in this recognition scenario. This evidence seems to indicate that MEG may be more appropriate for studying MCI, AD and other dementias when compared to EEG, though it may be obscured due to the fact MEG has been only recently used for this purpose.

It is also needed to objectively pointing out that the correct design of data analysis approaches is paramount to obtaining significant and reproducible results. For example in [58], the highest performance was 98%, which was obtained through a 5-fold cross-validation, whereas a performance of around 75% was also reported in the same study for the same data but using leave-one-out cross-validation, while using the same strategy of feature selection in both analysis. Due to the dramatic differences, we recommend to be cautious when interpreting such results and appreciations. In general, there are negative impacts and liabilities in using a relative small number of samples in these studies [28, 29]. Therefore, it is important for the M/EEG community to share their databases to ensure better reproducibility and enhanced statistical power by allowing multi-site statistical testing and data combination.

In addition, by further analysing the Table 3, a recent research by Amesquita-Sanchez et al. [46] proved to be quite promising. Different from the conventional biomarkers which derived from Fourier-based time-frequency analysis, the entropy-based biomarker employed in this work was built on top of empirical mode decomposition [68], a specially designed transform for non-stationary time-series. It is also noticed that in comparison to the performance of MCI detections, usage of MEG for AD detection still needs more investigations. Given the availability of mass AD databases in hospitals, a considerable boost to the research can be expected if more researchers will share/release the databases to the community.

3.3 Resting state comparison

In this section, we propose to compare some recent M/EEG system performances documented in the literature, while subjects remained in a resting state condition [69]. The resting state is one of the most common experimental scenarios and

Table 4. Resting state using EEG for MCI/AD detection, the sequence of the report is stacked downward based on the year of publication till the most recent

Feature(s)	Subjects	Channels	Recording	Epoch length	Performance
Relative logarithmic power spectral density values [57]	17 AD, 17 HC	128	3-4 minutes	N/A	Maximum accuracy 84.4%
Cross-frequency coupling measurements [38] (This is VEP task)	25 MCI, 15 HC	Pz	2 types of recordings each last for 30 seconds	1 second	Maximum accuracy 95%
Band power [21]	11 MCI, 11 HC	64	60 seconds	2.5 seconds	Theta relative power of MCI is larger than HC
Relative power in the conventional frequency bands, median frequency, individual alpha frequency, spectral entropy, Lempel-Ziv complexity, central tendency measure, sample entropy, fuzzy entropy, and auto-mutual information [44]	37 AD, 37 MCI, 37 HC	19	5 minutes	5 seconds	Maximum accuracy HC vs. All: 78.43% Maximum accuracy AD vs. All: 76.47%
Permutation dis-alignment index [20]	14 AD, 14 HC	16	10 minutes	8 seconds	Maximum accuracy 92.5%
Tsallis entropy, Higuchi Fractal Dimension, and Lempel-Ziv complexity [45]	32 AD, 20 HC	19	3 minutes	3 minutes	Maximum accuracy for single channel 85%
FFT-based power spectrum analysis, individual alpha frequency peak, transition frequency [62]	75 ADMCI, 75 PDMCI, 75 HC	19	5 minutes	2 seconds	Accuracy 63.48% ± 7.06%
Coherence measurements [60]	131 AD, 45 MCI, 45 HC	19	40 seconds	2 seconds	N/A

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

assumed as least affected by individual cognitive processes. The system performances reported based on this scenario are arguably comparable. Depending on the modalities (MEG and EEG), the most recent papers are listed in Table 4 and Table 5, respectively.

Table 4 lists a few recent papers reporting on MCI/AD detections using EEG data obtained in resting state. It is found that the power and entropy features are often considered as effective neuromarkers in detecting MCI and/or AD. Only a few minutes of recording appears to be enough for analysis, which is something worth pointing out. Usually in the clinical scenarios, there are ample recording length available. For MEG modality (Table 5), the entropy and complexity features are rather prevalent, whereas the power-based neuromarker appears to be less popular compared with EEG. The overall performance obtained using MEG for MCI/AD detection seems to be relatively worse than that of using EEG. A few suggestions can be made for better experiment design:

1) It seems most of the results are based on data obtained from single recording session, in contrast to longitudinal studies; such a scenario makes it quite difficult to address the issue of individual differences and expected changes associated with healthy ageing. Therefore, we suggest that research in MCI/AD changes should be combined with research in healthy ageing processes.
2) The recording length for each session should not necessarily be too long (a few minutes is found to be enough) in order to guarantee the quality of the features that capture the studied phenomena. Additionally, as AD affects multiple cognitive areas of the brain, diverse cognitive stimulation paradigms should be combined to better explore the cognitive changes.
3) Currently, the results are still incongruent, which can be attributed to the relatively small samples used in Alzheimer's research but also the disease heterogeneity, together with the wide range of used methodologies. For example, using the same database and features, the performances obtained through

Table 5 Resting state using MEG for MCI/AD detection, the sequence of the reports is stacked downward based on the year of publication till the most recent.

Feature(s)	Subjects	Channels	Recording	Epoch length	Performance
Lempel-Ziv complexity [40]	10 AD, 10 HC	148	5 minutes	20 seconds	ADs have less LZ complexity
Mean frequency, spectral entropy, approximate entropy, and Lempel-Ziv complexity [47]	18 AD, 18 HC	148	5 minutes	20 seconds	Without blind source separation (BSS): 72.2% of accuracy With BSS: 80.6% of accuracy
Shannon spectral entropy, Tsallis spectral entropy, generalized escort-Tsallis spectral entropy and Rényi spectral entropy, calculated from the relative spectral power [42]	20 AD, 21 HC	148	5 minutes	10 seconds	Maximum accuracy 85.4%
Sample entropy and multiscale entropy [41]	20 AD, 21 HC	148	5 minutes	10 seconds	Maximum accuracy 87.8%
Median frequency, spectral entropy, approximate entropy, Lempel-Ziv complexity [32]	20 AD, 21 HC	148	5 minutes	10 seconds	Maximum accuracy 80.5%
Coherence and synchronization likelihood [23]	18 MCI, 25 HC	148	5 minutes	24 seconds	Maximum accuracy 69.8%
Shannon spectral entropy, Tsallis spectral entropy, and Rényi spectral entropy (RSE), based on the normalized power spectral density [24]	18 MCI, 24 HC	148	5 minutes	10 seconds	Maximum accuracy of is 64.3% obtained from Shannon spectral entropy
Cross-approximate entropy [25]	12 AD, 12 HC	148	5 minutes	5 seconds	Maximum accuracy 70.83%
Disequilibrium: PSD-based Jensen's divergence [43]	36 AD, 18 MCI, 24 HC	148	5 minutes	5 seconds	AD significantly different from MCI; MCI slightly different from HC.
Structural connectivity [63]	22 sdMCI, 30 mdMCI, 29 HC	306	3 minutes	10 seconds	Maximum accuracy 86.27%
Functional connectivity metrics [70]	102 MCI, 82 HC	306	3-5 minutes	2 seconds	Maximum accuracy 83%
Dispersion entropy [26]	36 AD, 26 HC	148	5 minutes	10 seconds	*p*-value ranges from 0.006 to 0.114
Amyloid-beta deposition on regional power spectra [56]	28 MCI, 38 HC	306	5 minutes	10 seconds	N/A
Phase locking value, imaginary part and the correlation of the envelope [58]	24 MCI, 30 HC	306	4 minutes	10 seconds	Maximum accuracy 98% for 5-fold cross-validation; 75% for leave-one-out cross validation.

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

leave-one-out cross validation and n-fold cross validation have resulted in a change of more than 20% of accuracy [58]. In a real-life clinical scenario, current MCI/AD/HC classification accuracies are still not found sufficiently high as needed in clinical applications [58]. We exhort researchers to increase the sample size of ongoing studies and to share EEG/MEG database to openly comparing the developed methodologies.

IV. Discussion and Conclusion

This work explored the state-of-the-art neuromarkers from the ML perspective; its primary aim has been to review the mainstream algorithms devised for detecting/predicting MCI and AD using MEG and/or EEG signals. It is found that overall the MEG tends to perform slightly worse than EEG. However, as a relatively new modality, MEG has shown great potential in the field, particularly for MCI detection [46]. EEG with its much lower cost has been used for diagnosing the brain diseases for much longer time than MEG; the properties of the EEG signal are much well-studied and better understood. In general, there appears to be huge variability in prediction performance reported in the literature by algorithms using both EEG and MEG modalities. There is an urgent need to arrive at a consensus in terms of a single neuromarker or a neuromarkers combination most suitable for application in clinical practice.

Given its minor deployment in general, much work could be done in excavating the potential of MEG in MCI/AD detection. From the perspective of information gain, MEG modality intuitively should be much more informative than EEG signals, due to its higher sensor density and arguably more information per channel. Another strategy to boost the prediction performance is to combine the EEG and MEG: a multi-modal recognition system equipped with a well-designed fusion algorithm can synergistically combine complementary information from both modalities. From a broad perspective, using M/EEG based machine learning algorithms to classify AD patients and patients with other neurodegenerative diseases (such as Parkinson's disease, Lewy body dementia) may also deserve much attention in future investigations.

References

[1] "National life tables, UK: 2014 to 2016," *Office for National Statistics*, 2016. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014to2016. [Accessed: 30-May-2018].

[2] C. G. Gross, *Brain, vision, memory: Tales in the history of neuroscience*. MIT Press, 1999.

[3] "Dementia Fact sheet N°362," *World Health Organization, achived from the original*, 2015. [Online]. Available: https://web.archive.org/web/20150318030901/http://www.who.int/mediacentre/factsheets/fs362/en#. [Accessed: 30-May-2018].

[4] C. Ballard, *et al*, "Alzheimer's disease," *Lancet*, vol. 377, no. 9770, pp. 1019–1031, 2011.

[5] Derivative work: Garrondo (talk), "Alzheimer's disease brain comparison," *Wikimedia Commons*, 2008. [Online]. Available: https://commons.wikimedia.org/wiki/File:Alzheimer%27s_disease_brain_comparison.jpg.

[6] R. C. Petersen *et al.*, "Current concepts in mild cognitive impairment," *Arch. Neurol.*, vol. 58, pp. 1985–1992, 2001.

[7] "Alzheimer's Risk Factors & Causes | Alzheimer's Association." [Online]. Available: https://www.alz.org/alzheimers-dementia/what-is-alzheimers/risk-factors. [Accessed: 15-Aug-2018].

[8] A. J. Mitchell and M. Shiri-Feshki, "Rate of progression of mild cognitive impairment to dementia - Meta-analysis of 41 robust inception cohort studies," *Acta Psychiatr. Scand.*, vol. 119, no. 4, pp. 252–265, 2009.

[9] M. Längkvist *et al*, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognit. Lett.*, vol. 42, no. 1, pp. 11–24, 2014.

[10] S. Jirayucharoensak *et al*, "EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation," *Sci. World J.*, vol. 2014, 2014.

[11] J. M. Schoffelen and J. Gross, "Source connectivity analysis with MEG and EEG," *Hum. Brain Mapp.*, vol. 30, no. 6, pp. 1857–1865, 2009.

[12] E. Zamrini *et al.*, "Magnetoencephalography as a Putative Biomarker for Alzheimer's Disease," *Int. J. Alzheimers. Dis.*, vol. 2011, pp. 1–10, 2011.

[13] R. Hari *et al.*, "IFCN-endorsed practical guidelines for clinical magnetoencephalography (MEG)," *Clin. Neurophysiol.*, 2018.

[14] A. K. Andras Horvath *et al*, "EEG and ERP biomarkers of Alzheimer's disease: a critical review," *Front. Biosci.*, vol. 23, no. 1, p. 4587, 2018.

[15] A. Kumar *et al*, "Implication of preclinical biomarkers in identification of mild cognitive impairment (MCI) and Alzheimer ' s disease : Promises and future challenges," *Res. Reports*, pp. 1–16, 2018.

[16] Dubois, Bruno *et al*, "Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria.," *Lancet Neurol.*, vol. 13, no. 6, pp. 614–629, 2014.

[17] G. Q. Mamani *et al*, "EEG-based biomarkers on working memory tasks for early diagnosis of Alzheimer's disease and mild cognitive impairment," *2017 IEEE Healthc. Innov. Point Care Technol.*, pp. 237–240, 2017.

[18] A. Farooq *et al*, "Artificial intelligence based smart diagnosis of alzheimer's disease and mild cognitive impairment," in *2017 International Smart Cities Conference (ISC2)*, 2017, pp. 1–4.

[19] D. Domashenko *et al*, "Feature ranking for mild cognitive impairment and Alzheimer's disease diagnosis," in *2017 Signal Processing Symposium, SPSympo 2017*, 2017.

[20] H. Yu *et al*, "Functional brain connectivity in Alzheimer's disease: An EEG study based on permutation disalignment index," *Phys. A Stat. Mech. its Appl.*, vol. 506, pp. 1093–1103, 2018.

[21] J. Ye *et al*, "Relative Power Analysis of Electroencephalograph Rhythms in Elderly With Mild Cognitive Impairment," in *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017, no. 91420302, pp. 1–6.

[22] T. Tamura *et al*, "Magnetoencephalographic analysis of alpha frequency light entrainment of brain activity in dementia patients," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, 2003, pp. 2103–2105.

[23] G. Carlos *et al*, "Disturbed Beta Band Functional Connectivity in PatientsWith Mild Cognitive Impairment: An MEG Study," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 6, pp. 1683–1690, 2009.

[24] R. Bruna *et al*, "Analysis of spontaneous MEG activity in mild cognitive impairment and Alzheimer's disease using spectral entropies and statistical complexity measures," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.

[25] C. Gómez *et al.*, "Synchrony analysis of spontaneous MEG activity in Alzheimer's disease patients," *Conf Proc IEEE Eng Med Biol Soc*, pp. 6188–6191, 2012.

[26] H. Azami *et al*, "Dispersion entropy for the analysis of resting-state MEG regularity in Alzheimer's disease," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016, vol. 2016–Octob, pp. 6417–6420.

[27] M. S. Albert *et al.*, "The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's Dement.*, vol. 7, no. 3, pp. 270–279, May 2011.

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2019.2898871, IEEE Transactions on Biomedical Engineering

[28] P. Nguyen *et al*, "A Proposed Feature Extraction Method for EEG-based Person Identification," in *International Conference on Artificial Intelligence (ICAI). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).*, 2012.

[29] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017.

[30] G. Varoquaux, "Cross-validation failure: Small sample sizes lead to large error bars," *Neuroimage*, vol. 180, no. April 2017, pp. 68–77, 2018.

[31] G. Varoquaux *et al*, "Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines," *Neuroimage*, vol. 145, no. August 2015, pp. 166–179, 2017.

[32] R. Hornero *et al*, "Spectral and nonlinear analyses of MEG background activity in patients with Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 6, pp. 1658–1665, 2008.

[33] A. Mazaheri *et al*., "EEG oscillations during word processing predict MCI conversion to Alzheimer's disease," *NeuroImage Clin.*, vol. 17, no. October 2017, pp. 188–197, 2018.

[34] De Lange *et al*, "Interactions between posterior gamma and frontal alpha/beta oscillations during imagined actions," *Front. Hum. Neurosci.*, vol. 2, no. August, pp. 1–12, 2008.

[35] A. Mazaheri *et al*, "Functional Disconnection of Frontal Cortex and Visual Cortex in Attention-Deficit/Hyperactivity Disorder," *Biol. Psychiatry*, vol. 67, no. 7, pp. 617–623, 2010.

[36] Macmillan, Neil A., and C. Douglas Creelman, *Detection theory: A user's guide*, 2 nd. Psychology press, 2004.

[37] S. S. Poil, *et al*, "Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage," *Front. Aging Neurosci.*, vol. 5, no. OCT, pp. 1–12, 2013.

[38] S. I. Dimitriadis, *et al*, "A novel biomarker of amnestic MCI based on dynamic cross-frequency coupling patterns during cognitive brain responses," *Front. Neurosci.*, vol. 9, no. OCT, pp. 1–17, 2015.

[39] S. M. Pincus, "Irregularity and asynchrony in biologic network signals," in *Methods in enzymology*, 2000, pp. 149–182.

[40] G. Carlos, *et al*, "Decreased Lempel-Ziv complexity in Alzheimer ' s disease patients ' magnetoencephalograms," in *Engineering in Medicine and Biology 27th Annual Conference*, 2005, vol. 0, pp. 4514–4517.

[41] C. Gómez, *et al*, "Analysis of MEG recordings from Alzheimer's disease patients with sample and multiscale entropies," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2007, pp. 6183–6186.

[42] J. Poza, *et al*, "Analysis of Spontaneous MEG Activity in Patients with Alzheimer's Disease using Spectral Entropies," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 6179–6182.

[43] J. Poza, *et al*, "Analysis of spontaneous MEG activity in mild cognitive impairment and Alzheimer's disease using Jensen's divergence," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2014, pp. 1501–1504, 2014.

[44] S. Ruiz-Gómez *et al*, "Automated Multiclass Classification of Spontaneous EEG Activity in Alzheimer's Disease and Mild Cognitive Impairment," *Entropy*, vol. 20, no. 1, p. 35, 2018.

[45] Al-Nuaimi AH *et al*, "Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease," *Complexity*, vol. 2018, pp. 22–24, 2018.

[46] J. P. Amezquita-Sanchez *et al*, "A new methodology for automated diagnosis of mild cognitive impairment (MCI) using magnetoencephalography (MEG)," *Behav. Brain Res.*, vol. 305, pp. 174–180, 2016.

[47] J. Escudero *et al*, "Magnetoencephalogram Blind Source Separation and Component Selection Procedure to Improve the Diagnosis of Alzheimer's Disease Patients," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 5437–5440.

[48] Rasheed W. *et al*, "Moderate traumatic brain injury identification for MEG data using PU (Positive and Unseen) learning," 2017, pp. 75–78.

[49] C. J. Stam and B. W. Van Dijk, "Synchronization likelihood: An unbiased measure of generalized synchronization in multivariate data sets," *Phys. D Nonlinear Phenom.*, vol. 163, no. 3–4, pp. 236–251, 2002.

[50] M. Yu *et al*, "Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: An MEG-based multiplex network study," *Brain*, vol. 140, no. 5, pp. 1466–1485, 2017.

[51] A. S. Hincapié *et al*., "The impact of MEG source reconstruction method on source-space connectivity estimation: A comparison between minimum-norm solution and beamforming," *Neuroimage*, vol. 156, no. May, pp. 29–42, 2017.

[52] A. Fuchs, "Beamforming and its applications to brain connectivity," in *Understanding Complex Systems*, vol. 2007, 2007, pp. 357–378.

[53] "CTF MEG, Canada." [Online]. Available: https://www.ctf.com/. [Accessed: 27-Nov-2018].

[54] B. D. Van Veen *et al*, "Localization of {B}rain {E}lectrical {A}ctivity via {L}inearly {C}onstrained {M}inimum {V}ariance {S}patial {F}iltering," *Ieee Trans. Biomed. Eng.*, vol. 44, NO. 9, no. 9, pp. 867–880, 1997.

[55] L. Douw *et al*, "Consistency of magnetoencephalographic functional connectivity and network reconstruction using a template versus native MRI for co-registration," *Hum. Brain Mapp.*, vol. 39, no. 1, pp. 104–119, 2018.

[56] A. Nakamura *et al.*, "Electromagnetic signatures of the preclinical and prodromal stages of Alzheimer's disease," *Brain*, no. March, 2018.

[57] H. Aghajani *et al*, "Diagnosis of early Alzheimer's disease based on EEG source localization and a standardized realistic head model," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 6, pp. 1039–1045, 2013.

[58] S. I. Dimitriadis *et al.*, "How to build a functional connectomic biomarker for mild cognitive impairment from source reconstructed MEG Resting-state activity: The combination of ROI representation and connectivity estimator matters," *Front. Neurosci.*, vol. 12, no. Jun, pp. 1–21, 2018.

[59] J. Guillon *et al*, "Loss of brain inter-frequency hubs in Alzheimer's disease," *Sci. Rep.*, vol. 7, no. 1, pp. 1–14, 2017.

[60] Medvedeva, A. V., and N. N. Yahno, "Functional Connectivity as a Neurophysiological Biomarker of Alzheimer's Disease," *J. Alzheimer's Park. Dement.*, vol. 3, no. 1, pp. 1–8, 2018.

[61] Chatrian, G. E. *et al*, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity," *Am J Electroencephalogr Technol*, vol. 25, pp. 83–92, 1985.

[62] C. Babiloni *et al*, "Functional cortical source connectivity of resting state electroencephalographic alpha rhythms shows similar abnormalities in patients with mild cognitive impairment due to Alzheimer's and Parkinson's diseases," *Clin. Neurophysiol.*, vol. 129, no. 4, pp. 766–782, 2018.

[63] J. A. Pineda-Pardo *et al.*, "Guiding functional connectivity estimation by structural connectivity in MEG: An application to discrimination of conditions of mild cognitive impairment," *Neuroimage*, vol. 101, pp. 765–777, 2014.

[64] L. F. Haas, "Hans berger (1873–1941), richard caton (1842–1926), and electroencephalography," *J. Neurol. Neurosurg. Psychiatry*, vol. 74, no. 1, pp. 9–9, 2003.

[65] E. A. Curran and M. J. Stokes, "Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems," *Brain Cogn.*, vol. 51, no. 3, pp. 326–336, 2003.

[66] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cogn. Emot.*, vol. 23, no. 2, pp. 209–237, 2009.

[67] S. Yang and F. Deravi, "On the Usability of Electroencephalographic Signals for Biometric Recognition : A Survey," *IEEE Trans. Human-Machine Syst.*, vol. 47, no. 6, pp. 958–969, 2017.

[68] N. E. Huang *et al*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[69] Engels, M. M. A. *et al*, "Alzheimer's disease: The state of the art in resting-state magnetoencephalography," *Clin. Neurophysiol.*, vol. 128, no. 8, pp. 1426–1437, 2017.

[70] Maestú, F. *et al*, "A multicenter study of the early detection of synaptic dysfunction in Mild Cognitive Impairment using Magnetoencephalography-derived functional connectivity," *NeuroImage Clin.*, vol. 9, pp. 103–109, 2015.

0018-9294 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.